(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

C12N

Jose, CA 95117 (US). DRMANAC, Radoje, T. [YU/US];

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09/496,914 09/560,875 3 February 2000 (03.02.2000)

27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US Filed on

Filed on

09/496,914 (CIP) 3 February 2000 (03.02.2000)

US

09/560,875 (CIP) 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston,

- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
- HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

MA 02111 (US).

without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) Abstract: The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

DAICDOOID JAKO

045740040 | .

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

			•	1 'Y
				A Transfer
•		•		
,				
				•
			•	
	•			

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	110.			Score	100.101
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68 .
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	375	43
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein (hs6M1-21))	480	55
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410_5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KJAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- 57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

		•	
·			

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	Species	Description	Waterman	Identity
NO:	140.			Score	Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
902	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
903	Y00261	Homo sapiens	Human secreted protein, SEQ ID NO: 3430. Human secreted protein encoded by gene 4.	1133	199
				771	99
905	AF039688	Homo sapiens	antigen NY-CO-3	2544	100
906	AB007836	Homo sapiens	Hic-5		
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo 1	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162 1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP- 62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
921	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member	2763	99
021	V/22220		GLUT9		100
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924 G03827	Homo sapiens	B-cell receptor associated protein Human secreted protein, SEQ ID NO: 7908.	1469 529	100
934 935	AB039371	Homo sapiens Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
930	V20292	familiaris	rabe	1004	100
937	B08906	Homo sapiens	Human secreted protein sequence encoded by	117	44
020	N (12 (00	77	gene 16 SEQ ID NO:63.	1000	100
938 939	M13692 Y53886	Homo sapiens Homo sapiens	alpha-1 acid glycoprotein precursor A suppressor of cytokine signalling protein	1064 515	99 42
940	Y16630	Homo sapiens	designated HSCOP-6. Human Putative Adrenomedullin Receptor	1904	99
941	AC005102	Homo sapiens	(PAR). small inducible cytokine subfamily A member	627	99
	1	ļ. <u>. </u>	24	L	
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
		Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

			••
		•	
		,	
		-	

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
1051	W78324	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (cione ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	262	100
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065 1066	Y41674 AJ250042	Homo sapiens Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066	Y36087	Homo sapiens	Rab5 GDP/GTP exchange factor homologue Extended human secreted protein sequence, SEQ	2575 770	100 85
1068	Y94959	Home sapiens	ID NO. 472. Human secreted protein clone mc300 1 protein	301	100
1069	Y94959	Homo sapiens	sequence SEQ ID NO:124. Human secreted protein clone mc300_1 protein	301	
1070	W64535	I	sequence SEQ ID NO:124.		100
1071	X03145	Homo sapiens Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014 148	99
1072	AL031177	Homo sapiens	pot. ORF III dJ889M15.3 (novel protein)	821	50 91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	35
1076	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899	Homo sapiens	ribosomal protein L10	332	76
1078	Y91447	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080	AB039723	Homo sapiens	WNT receptor frizzled-3	1376	92
1081	AB020527	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1082	L13802	Homo sapiens	ribosmal protein small subunit	499	80
1083	W75098	Homo sapiens	Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1085 1086	G04063 AF090942	Homo sapiens	Human secreted protein, SEQ ID NO: 8144. PRO0657	88	43
1087	G00517	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	124 129	64
1088	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 4396. Human secreted protein, SEQ ID NO: 8172.	126	36
1089	AF140631	Homo sapiens	G-protein coupled receptor 14	364	82
1090	G04063	Homo sapiens	Human secreted protein, SBQ ID NO: 8144.	114	32
1091	572304	Mus sp.	LMW G-protein	146	83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone HEMAM41.	405	100
1093	W85612	Homo sapiens	Secreted protein clone fh123_5.	4358	97
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein sequence SEQ ID NO:30.	1013	99
1095	Y92345	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-62.	409	100
1096	AF090942	Homo sapiens	PRO0657	147	60
1097	L24521	Homo sapiens	transformation-related protein	166	58
1098	X56932	Homo sapiens	23 kD highly basic protein	490	70
1099	G04063	Homo sapiens	Human secreted protein, SBQ ID NO: 8144.	83	35
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	149	59

rake trapel ik Mendelik di kecamatan di kecamatan di kecamatan di kecamatan di kecamatan di kecamatan di kecam

. . . $|x_{ij}| = |x_{ij}| + |x_{ij}|$

070 10	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID			1 -			D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	{	i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l	İ	1.	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		1		Sequence	/=possible nucleotide deletion. \=possible
ţ	1	į.	('	peptide	{	
	<u> </u>			sequence		nucleotide insertion
	1	1			}	VTLRVTGESHIGGVLLKIVEQINRKQDWSDH
	l	1	1	Į.	ļ	AIWWEQKRQWLLQTHWTLDKYGILADARLF
i	j	l	1		Ī	FGPQHRPVILRLPNRRALRLX*
658	2008	Ā	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLQPPPPGF
""	2005	1	1 2027	•		KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
						HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
	0000	 	-		220	
659	2009	A	5018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
		i	1			T*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
	1					MAACWAVHVKTHMRPGLAVLPRLVLNSWS
1	ı	Í	Í		ĺ	*AIILLWPPKALGLQA
660	2010	A	5028	2	310	SRVDDFVGERRGGCDECLCGHRGLRAVPLG
1	1	1				HPGHLCLOPPGGPA*FLDYCRGCCPHPVPGST
1	1	1]		1	AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH
i	i	1 .	1)	}	HRTGARWNH
-	+ <u></u>	 	+ ====	I	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
661	2011	Α	5050	752	431	
1	j	1	1	1	1	LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS
	1	1			•	LELLGSSHPPTSASQSARITGVSHRAWPLK*F
}	}	Ì	ł		l	NLNQYQTLTMN
662	2012	A	5054	48	103	ELNNGPFQMPLCNGGNLAVTGSWADRSPLH
1					t	EAASQGRLLALRTLLSQGYNVNAVTLDHVTP
						LHEACLGDHVACARTLLEAGANVNAITIDGV
}	}	}	1	}	ļ	TPLFNACSQGSPSCAELLLEYGAQAQLESCLP
1				l		SPTHEGASKGHHECLDILISWGIDVDQEIPHSG
1		1		ļ		TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
Į.	1	1	Į.	1	Į.	
						WDTPLPGAGHQSTQKLE*LFAMVEIWQ
663	2013	Α	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK
1	İ	[1	1	Ì	ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
i			1			GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
l	•	1	ł	i		WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
664	2014	Α	5071	550	1	LSFIEVLSMEQVNKTVVREFVVLGFSSLARLQ
•••		1			_	OLLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP
		1	1			MYFFLAILSCSEICYTFVIVPKMLVDLLSOKK
į	Į.	1.	ł	ł i	ł	TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
	į.					YMAICNPLRYSVLMGHGVCMGLMAAAWAC
1	ļ	1				
	<u> </u>			<u></u>		GFTVSLVTTSLVFHLPFHSSNQHE
665	2015	A	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
•		1			1	PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
1			ŀ	1	1	HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD
	1	1	1	1		HDQNEGFHCREECRILGHSDRCWMPRNPMPI
	1	1	1	1		RSKSPEHVRNIIALSIEATAADVEAYDDCGPT
]	I]	J	1	J	KRTFATFGKDVSDHPAEERPTLKGKRTVDVT
ì	1	1	Ì	1	1	ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL
}	1	1	1	ŀ		
		<u> </u>	 	L	L	PTKPSVSYEIVDPGITARRC
666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
ſ	1	1	1			VGPKAQLSGLQLQPCLYKRR
667	2017	Α	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
1	1	1	1	}	l	PCVY*RR
668	2018	A	5086	852	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
000	2018	^	3000	032	233	RALPTTFADIENLKYLLFTRDASQPFYLGHTV
[1	1	1	1	1	
-	1	1	İ			IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE
]]	J	1]		TCADQSVIWKLSEDKQLAICLKYAGVHAENA
1	1	1	1	Į.	ļ	EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG
	1	1		1		CCSDMAITFNGLTPQKMEVMMYGLYRLRAF
1	1	ı	1			GHYFNDTLVFLPPVGSEND
669	2019	A	5101	1	329	PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG
003	2019	1^	1 2101	1	329	
1	1	1	1			ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP
1	1	1	1		1	RGCQHEAAPCPRGPGSDGLHHASAACASLPP
<u>_</u>				l	1	SPILPVLLPELGPL
670	2020	A	5102	3	547	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP
						·

	•
-	
•	

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		Í		peptide sequence	•	nucleotide insertion
<u> </u>				sequence	 	PKELVRKPYVLNDLEAEASLPEKKGNTLSRD
	<u> </u>	}	İ		1	LIDYVRYMVENHGEDYKAMARDEKNYYOD
ľ		l			1	TPKQIRSKINVYKRFYPAEWQDFLDSLQKRK
]				MEVE
932	2282	A	8011	412	1	SNLCLGNSWRWRWAKSRHHCIPTVTLSKRSG
}	ļ -	i i]	İ	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
						QGRGQIPIPCPWPPPPPPPPPGSPGPGCRQFHQ
		ŀ				SLEAKARHPASVREMRGKVKMRRALRRAPA
						STRASSRQPNPK
933	2283	Α	8012	147	1077	PPVPPASRSDMAQNLKDLAGRLPAGPRGMGT
	}				Ì	ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF
1						NRIGGVQQDTILAEGLHFRIPWFQYPIIYDIRA
		ł			!	RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
!		•	!			PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF
ļ						NASQLITQRAQVSLLIRRELTERAKDFSLILDD VAITELSFSREYTAAVEAKQVAQOEAORAOF
1				ļ		LVEKAKQEOROKIVQAEGEAEAAKMLGEAL
		}		}	J	SKNPGYIKLRKIRAAQNISKTIATSONRIYLTA
						DNLVLNLQDESFTRGSDSLIKGKK
934 ·	2284	A	8023	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ
1		**	0020			RLRKFRELHLMRNEARKLNHOEVVEEDKRL
1) 	ł			ì	KLPANWEAKKARLEWELKEEEKKKECAARG
						EDYEKVKLLEISAEDAERWERKKKRKNPDLG
	<u> </u>					FSDYAAAQLRQYHRLTKQIKPDMETYERLRE
}	}	ł	1]	}	KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE
ļ						KQIEKRDKYSRRRPYNDDADIDYINERNAKF
	0005	ļ.,	0000		210	NKKAERFYGKYTAEIKQNLERGTAV
935	2285	Α	8027	59	310	LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL
						QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL SQHSSPAPMYSQTFHILVLG
936	2286	A	8032	1	639	SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL
750	2200	11	0032	i •	037	FRESEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ
1	ŀ		[IWDTAGQERFRTITTAYYRGAMGIMLVYDIT
	Ì					NEKSFDNIRNWIRNIEEHASADVEKMILGNKC
			ļ			DVNDKRQVSKERGEKLALDYGIKFMETSAK
l		l	1			ANINVENAFFTLARDIKAKMDKKLEGNSPQG
L	<u> </u>	L				SNQGVKITPDQQKRSSFFRCVLL
937	2287	Α	8039	393	311	EETIHSENSYILEKYIPISANLTLTIA
938	2288	Α	8052	675	.1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLL
						PLLEAQIPLCANLVPVPITNATLDRITGKWFYI
	1					ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF
		1				LREYQTRQDQCIYNTTYLNVQRENGTISRYV
]					GGQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYEALDCLRIPKSD
	1	l	1	1		VVYTDWKKDKCEPLEKQHEKERKQEEGES
939	2289	A	8055	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF
""	2207	1 ^	5055	1	1000	AEQLKWSAELARLGESIMDGKQGGMDGSKP
}]]])	AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
						IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
1						CLGGHLSCVKILLKHGAQVNGVTADWHTPL
		ŀ	1		[FNACVSGSWDCVNLLLQHGASVQPESDLASP
	1	1	ſ	1		IHEAARRGHVECVNSLIAYGGNIDHKISHLGT
		1				PLYLACENQQRACVKKLLESGADVNQGKGQ
						DSPLHAVARTASEELACLLMDFGADTQAKN
]	[1	1	ł	l	AEGKRPVELVPPESPLAQLFLEREGPPSLMQL
	1			_		CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH
			00.55	<u> </u>	1	L
940	2290	Α	8058	2	1203	KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI
Г.	L	<u> </u>	L	<u> </u>	<u> </u>	ARLRTMDLHLFDYSEPGNFSDISWPCNSSDCI

			• .
	٠.		•
		•	
			·
•			
-			
			•

	SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
Desired Sequence	4		noa			,	
Sequence			l				
1073			1)			
minio acid residue of peptide requence T-Threonine, V=Valine, W=Typtoplan, Pypeside bequence Peptide requence Peptide r	1 7	delice	ł				
Peptide Pept	1	ĺ	1	-1.			
peptide	ł			1			
	1	1	i	1		}	
CYSHQMAQMGVGEOKSIGEWVLGPNTVQSQQVKNLALIPEWNSLGJVYVSMDNPSGSIA 1073 2423 A	ļ	j		j	sequence		nucleotide insertion
1073 2423 A 8879 146 412 DFSV*GDVDEVTCPLCQLITEPISLINGCIRL							CYSIHQMAQMGVGEGKSIGEWVLGPNTV\AQ
1073	i			ì			
QVCITA **RESYUSGG** **SSSPVCHTTOPANIL RISRYLP* **SLIGPIDPIGG** **STRYLP*	1						
1074	1073	2423	Α	8879	146	412	
1074	1		1		<u> </u>		
REISFGDYICHTFOODCWADRSPILEAAAHO RLIALKTILAQGVIVNILDRYSSLHEACL CPVACAKPYWKWPRHOGTVTGPPLLMY CPVACAKPYWKWPRHOGTVTGPPLLMY CPVACAKPYWKWPRHOGTVTGPPLMY CPVACAKPYWKWPRHOGTVTGPPLMY CPVACAKPYWKWPRHOGTVTGPPLMY CPVACAKPYWKWPRHOGTVTGPPLMY CPVACAKPYWKWPRHOGTVTGPPLMY CPVACAKPYWKWPRHOGDENITH CPVACAKPYWKWPRHOGDENITH CPVACAKPYWKORKISKERINGLOAPP VWGLSPKNOFTPCSDENITH VWGLSPKNOFT VWGLS				<u> </u>		<u> </u>	
RILALKTILAQGYNVNILWIT.DRYSSI.HEACI	1074	2424	Α	8884	67	435	
1075 2425 A 8896 1294 248 RSGDRIGTHQLGGLSGSRINGSYRSSRSSRS RSERPSAPRGIPPASASSSVYJGSVRPVGSDR PWPSLLDKERESIR ROSELGAPE VWGLSPKNPEDPSDEHTTPVEDEEPKKSTTSAS TSEBEKKKSSKKERKKSKKKSTKKSTKKSTKAKK KYSEDSDADEDSETTSDEDNKRRAKKAKKK KKKKKRSKKYKKKSKSKKIKK KYSEDSDADEDSETTSDEDNKRRAKKAKKK EKKKKRRSKKYKKKSKSKKIKK KKKKRRSKKYKKKSKSKKIKK KKKKRRSKKYKKKSKSKKIKK KKKKRRSKKYKKKSSKKIKK KKKKRRSKKYKKKSKSKKIKK KKKKRRSKKYKKKSKSKKIKK EKKKKRRSKKYKKKSKSKKIKK EKKKKRRSKKYKKKSSKKIKK EKKKKRRSKKYKKKSSKKIKK KKSEDSDADEDSETTDSDDEDNKRRAKKAKKK EKKKKRRSKKYKKKSSKKSKIKESDSSSSES GEETLENPWKDETKANQVM**VYSSHER MEAVRTAKEEPSTVLMREPLIPPNPRRET ERRE SRETE KERE SRETE KERE SRETE KERE STYNSSREVKABSEPDLIGPEAPKTLTS QDDKFLTAKKGSKRFSKPFSGKS FYLDLPAGKNLQFLTGAIQQLGGVEGFLSKE SYSYSSREVKABSKRFSKPSPSGKS FYLDLPAGKNLQFLTGAIQQLGGVEGFLSKE SYSYSSREVKABSKRFSKPSPSEVR VETSAMVDIKGSHPRPSKRPVDSYPLSRGKE LIQKAIRQK**GTVOSHCHLAVKGV VYPALHRAGTEWQLSALHRAPRSTOPDKAC RSGREHVQQQSGBHGKWPDLKGPR KTANGKAQVIVIPRCKGOWKCYPYKAVT VGKPVHHIGVNPLKFAQSLOSVABEQ A 8901 352 3 AKGAYKYIQELWKRYGGWKKYPOKKAT VYGKPVHHIGVNPLKFAQSLOSVABEQ A KROSE KAKKGAQVIVIPRCKGOWKCYPYKAVT VYGKPVHHIGVNPLKFAQSLOSVABEQ CRISCEPTPLAKQQAYLPSKCHGOWKCYPYKAVT VYGKPVHHIGVNPLKFAQSLOSVABEQ CRISCEPTPLAKQQAYLPSKCHGOWKCYPYKAVT VYGKPVHHIGVNPLKFAQSLOSVABEQ CRISCEPTPLAKQQAYLPSKCHGHLHGPLH RP*QAAPALHAGQQLAPHIPPL RP*QAAPALHAGQQLAPHIPPL RP*QAAPALHAGQQLAPHIPPL RP*QAAPALHAGQQLAPHIPPL CYPKLI STYNKNGHLSFKVVLUDNYDLASSPVEANKYI CRISCEPTPLIKQOPLYNTTIFMPLLNHY DQQTAALAMEPPHIPMVNLOCSRFFFTLCAL A A S920 381 1788 SSESPSDPGRMAMTWIVPSKYGCSRFFTLCAL A A SPICE STANYFIGILEDR VACANSIPAQVKA STLYTGGSINKACTIM STYNKNGHLSFKVVLUDNYDLASSPVEANKYI CRISCEPTILIKAQDSINGVCYGCYGLXD VYDALRYFVLAPLCLYVVGYSLLLAGIISLNR VRIEDPL**REGORDISGCYCYGLXD VYDALRYFVLAPLCLYVVGYSLLLAGIISLNR VRIEDPL**REGORDISGCYCYGLXD VYDALRYFVLAPLCLYVVGYSLLLAGIISLNR VRIEDPL**REGORDISGCYCYGLXD VYDALRYFVLAPLCLYVVGYSLLLAGIISLNR VRIEDPL**REGORDISGCYCYGLXD VY				j	1	1	
1075	İ	i	1	Ì	ſ		
SRERPSAPRGIPPASASSSVYYGSVSRYYGSDK PWPSILDKERSERICKELGAPE	1000	0.00	ļ.,	0000	1004		
PWPSILDKEREESI.RQKRI.SERRIGELGAPE VWGLSPKNPEPDSDEDDMERRAKSTSAS TSPIECKKKKSSRSKERSKKRKKKSTSAS TSPIECKKKKSSRSKERSKKRKKKSSTSAS TSPIECKKKKSSRSKERSKKRKKKSSKRKHK KYSEDSDEDDSDDMERRAKAKKAKKKSKKKKKKSKKKKKSKKKKSKKKKS QBEFLENPWKDDTKABEPSDLDMKRAKAKKK EKKKKHRSKYYKKKRSKKRKESSDSSSKES QDDKPLINYOHALLPGGAAMAEYYKAGKI PRRGEIGLTR*RNCHHLINAQVM**VYSRHRR MEAVRTAKREPESTPIPPFRETE KERE GRSTEAEKEPAFDERTGKGRRI.PRAGEPHG*E *APGFGPRSFQVSRKMPESPPGARKHPPSGKS VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VETSAMVDFKGSILPRAGUPLGVIEGISKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VETSAMVDFKGSILPRAGUPLGVIEGISKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VETSAMVDFKGSILPRAGUPLGVIEGISKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VETSAMVDFKGSILPRAGUPLGVGSTSKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VSTAMVDFKGSILPRAGUPLGVGSTSKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VSTAMVDFKGSILPRAGUPLGVGSTSKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VETSAMVDOGNGSLEVPDLLGGFT.SKE VSYIVSSRRBVKABSSGKSIRGCPSFSPSEVR VSTAMVDFKGSILPRAGUPLGVGSFSFSEVR VSTAMVDFKGSILPRAGUPLGVGSFSPSEVR VSKRJARHVOQNSQEMDLKGPSISKPVDLKGPTSLKC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPDKAC QYPALHRAGTLBWQLSALHRAPRSTQPBKAC PRESENDAL PRAGUPLACH PROBLEM PROBLE	10/5	2425	A	8820	1294	248	
VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS TSEERKKKKSSRKERSSSSKES TSEEKKKKSSRSKERKSKKKSSSKKEKSSSSKES QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGGBAMAEYVKAGKRI PREGEIGLTR*PKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGGBAMAEYVKAGKRI PREGEIGLTR*PKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGGBAMAEYVKAGKRI PREGEIGLTR*PKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGGBAMAEYVKAGKRI PREGEIGLTR*PKDR*VYSRHRR MEAVRTAKREFESTVLMRREFLHPFNPRRET KERE REVER A 8899 146 789 GRSTEAEKEPAPDERTGKGRRLPFAGEFHG*PE APGPEPRSFQVSRKMPESUPGARKHPPSGKS FYLDLPAGKNLQPLTGAIQQLGGVBGFI SKE VSYTVSSREVKAESSGKSHRGCPSPSSEVR VSTSAMVDFKGSIRTFSRKFVDSVPLSRGKE LLQKAIRNQK**CTVQQLSHCRLYGGEKTTAK RSQREHVQQOSQEHGKWPDLKGFR VSTVSSREVVKAESSGKSHRGCPSPSSEVR RLQYKAKQQYIVTRICVRRGGWKCPVPKAVT VGKPVHHGVN*LKAGSLGVMHFLLRVRCW QYPALJRRAGTEWQLACRGGRRRGHLHGPLH RP*QAAPALHAGQQLAPHAFPT 1078 2428 A 8905 536 781 ACPAENEEVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRGHLHGPLH RP*QAAPALHAGQQLAPHAFPT 1079 2429 A 8912 121 376 NLIWKLLCVTERRLVILDNYDLASEYVEANKYI CRRIIGPKRGQDLACRGGRRRGHLHGPLH RP*QAAPALHAGQQLAPHAFPT VGKPYHGMV*LTGLAFTGATTACHTAGAGVKYI CRRIIGPKRGQDLACRGGRRRGHLHGPLH RP*QAAPALHAGQQLAPHAFPT VTKLGHTSTFL*CYTKLI EYNKNGHLSFKYVKTTSMDEY SSESPSDFGRMAMTWIVFSLWPLTVFMGHIG GHSLPSCEPITLRMCQDLPYNTTFMPNLLNHY DQTAALAMEFPHMVNLDCSRDFFFTLCAL YAPICMEYGRVTLPCRRLCQDRAYSECSKLME MGPGTYWPPDMECSPPCDCEPYPTRLVLULNA GEFTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCDCHPYPTRLVLULNA GEFTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCDCHPYPTRLVLULNA GREFTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCDCHPYPTRLVLULNA GREFTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPRFTLCALL YAPICMEYGRVTLPCRRLCQDRAYSECSKLME WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKWGSEAIEKKALLFHASA WGREFTLTHLAAVPKTACTORISTCHACGUCPWGTSLLVLVELL VVIGCYFYSOAYRGIWSITWQEEC CSKKGIVDEFFTLLS		1				į	
TSEEEKKKKSSRSKERSKKRKKSSKKKHK KYSEDSDSDSDSDSDEDNIKRAKKAKKK KYKKHRSKKYKKKSKKSRKESSDSSSKES QEFLENPWKDRITKAEEPSSDLIGPRAPKILTS QDDKPLNYGHALLPGEGAAMAEYVKAGKRI PRRGEIGLTR-RNCHLINAQVM**VVSRHRR MEAVRIAKEPEESTVLMRRELIPFNPRRET KERE	ł	ļ	(1	(Į.	
KYSEDSDSDSETDSSDEDNKRRAKKAKKK EKKKKHRSKYYKKRSKSRKSPSSSSSSSSSS OBEFLENPWKDRTKAEEPSDLIGPEAPKTLTS QDDKPLNYGHALLPGBGAAMAEYVKAGKI PRRGEIGLTR*RNCHILNAQVM**VVSRHRR MEAVRTAKREFESTVLMKREFLHPFNPKRET KERE	1	ļ]				
			1				
QEEPLENPWKDRTKAEEPSDLIGPEAPKTLTES QDDKPLNYGHALLPGEGAAMAEYVKAGKRI PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR MEAVRTAKREPESTVLMRREPLHFFNPRRET KERE		ļ	ĺ	i			
	Í	1	ſ	1	ĺ		· · · · · · · · · · · · · · · · · · ·
PRRGEIGLTR*RNCHHLNAQVM***VYSRHRR MEAVRTAKREPESTVLMRREPLHPFNPRRET KERE	i	i	į	İ	į	į	
	1		1	j	Ì		
		į	1	Į.	[İ	MEAVRTAKREPESTVLMRREPLHPFNPRRET
*APGPGPRSFQVSRKMPEEIPPGARKIIPFSGKS FYIDLPAGKNLQFITGAIQQLGGVIEGFLSKE VSTYINSSREVKAESSGKSHRGCPSPSPSEVR VSTSAMVDPKGSIPRPSRKPVDSVPLSRGKE LLQKAIRNQK**CTIVQQLSHCRLVGEKTTAK RSQREHVQQQSQBHGKWPDLKGPR RSQREHVQQQSQBHGKWPDLKGPR AKIGAYKYIQELWRKKQSDVMHFLLRVRCW QYPALHRAGTEWQLSALHRAPRSTOPDKAC RLGYKAKQQYIIYRICVRRGGWKCPVPKAVT YYGKPVHHGVV*LKFAQSLQSVAEEQ ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRQHI.HGPLH RP*QAAPALPGSRQALCRGGRRQHI.HGPLH RP*QAPALHGQCLAPHPPT RP*QAPALHGQCLAPHPPT CNIIQFKPGQDKYFILGLPTOSTFL*CYPKLI EYNKNGHI SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLPSCEPITLRMCQDLPYNTTJMPNLLNHY DQQTAALAMEPFHPWNLLDCSRDFRPFLCAL YAPICMEYGRVTLDCRACQAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPCPFMYFRREELSFARVFIGLIS IICLSATLFIFVTFILDVTRFYPERPIKCYAV WHMMVSLIFFVGJLEDRVACNASIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAVPK WGSEAIEKKALLFHASA WGIPGTLTILLLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWSTTWIQERC CSRKGIVDEFFPLLALLAGISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWSTTWIQERC CSRKGIVDEFFPLLSN*CIWTOPQGYPQSSYG TLANPVF\CSVRHGILALLLQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST CANADAT CANA		1	1				
FYLDLPAGKN_QFLTGAIQQLGGVEGFLSKE VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR VSTSAMVDFKGSHPRSRKPVDSVPLSRGKE LLQKARNQK**CTVQQLSHCRLJVGEKTTAK RSQREHVQQQSQEHGK WPDLKGPR VSYALHRAGTEWQLSALHRAPRSTQPDKAC RIGKAKQGYIVTRICVRRGGWKCPVPKAVT VGKPVHHGVN*LKFAQSLQSVAEEQ VSYALHRAGTEWQLSALHRAPRSTQPDKAC RLQYKAKQGYIVTRICVRRGGWKCPVPKAVT VGKPVHHGVN*LKFAQSLQSVAEEQ VSYALHRAGTERQHLHGPLH RP*QAAPALHAGQQLAPHAGFQAGPQGPA PALPFAATPGSRQAALCRGGRRQHLHGPLH RP*QAAPALHAGQQLAPHPPT VIIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIQFRPGQDKYPTLGPTPGSTEL*CYPKLI EYNKNGHLSFKVVKTFSMDEY SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLPSCEPTILRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLQCAYSECSKLME MFGVPWEDMECSRFFDCDEFYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SPLLVRDCSPPCPNMYFREELSFARYFIGLIS IICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFLGFLLEDRVACNASIPAQYKA STLVTQGSINKACTMLFMILYFFTMAGSVWA WHMMVSLIFFLGFLLEDRVACNASIPAQYKA STLVTQGSINKACTMLFMILYFTMAGSVWA WHMMVSLIFFLGFLLEDRVACNASIPAQYKA STLVTQGSINKACTMLFMILYFTMAGSVWA WHMTVFLAAVFKWGSEAIEKKALLHASA WGIPGTLTITUFLAAVFKWGSEAIEKKALLHASA R VRIEIPL*KENQDKLVKFMRIGYFSILTYLVPLL VVIGCYFYEQAYRGIWETTWIQERC CSRKGIVDEFFPLLSN*CIWTQPQQFPQSSYG TLANFVFCSVRHGLALILQLCNFISIYTQQMN LSIAIFAAMVNNTAPPSQPNASTERPST	1076	2426	A	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
1077 2427 A 8901 352 3 AKIGAYKYIGELWRKKQSDVMIFLRVRCW RSQREHVQQQSQEHGK WPDLKGPR 1077 2427 A 8901 352 3 AKIGAYKYIGELWRKKQSDVMIFLLRVRCW QYPALHRAGTEWQLSALHRAPRSTQPDKAC RLGYKAKQGYITYRICVRGGWKCPVPKAVT YGKPVHIGVN*LKFAQSLQSVAEEQ 1078 2428 A 8905 536 781 ACPAENREVPEMAAGQAPHAGFGAGPGQFA PALPFAATPGSRGQALCRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASSVEANKYI CHRIGOFFAQDOKYPTLGLPTGSTEL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDFGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPHLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVAVQVQDYGFWCPRELKIDPDLGY SPLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFFVFTLDVTRFYPERPIKCYAV WHMMVSLIFFVIGFLLEDRVACNASIPAQYKA STVTQGSHNKACTMLFMILYFTMAGSVWW VILITTWFLAVPKWGSEAIEKKALLFHASA WGIPGTLTILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMRIGVFSILYLVPLL VVIGCYPQAYRGIWEITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWIOPQGYPQSSYG TLANFVFCSVRHGLALILQLCNFISYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	ł	ł	ł	! .	i	1	
VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RSQREHVQQVSQEHGKWPDLKGPR RRQYKAKQGYIIYRICVRRGGWKCPVPKAVT YGKPVHHGVN*LKFAQSLQSVABEQ RLQYKAKQGYIIYRICVRRGGWKCPVPKAVT YGKPVHHGVN*LKFAQSLQSVABEQ RQFVKAKQGYIIYRICVRGGWRCPVPKAVT YGKPVHHGVN*LKFAQSLQSVABEQ RAPAALHAGQQLAPHAPT RP*QAAPALHAGQQLAPHAPT RP*QAAPALHAGQQLAPHAPT RP*QAAPALHAGQQLAPHAPT RP*QAAPALHAGQQLAPHAPT RP*QAAPALHAGQQLAPHAPT TORIO ROMANI RYKNCHLSPKYVKTFSMDEY SSESPSDFGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GBETEGAPVAVQRDYGFWCPRELKIDPDLGY SSLHVRDCSSPCFPMYFRREBLSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFVGFLEDVACNASIPAQYKA WHMMVSLIFFVGFLEDVACNASIPAQYKA STVTQGSHNKACTMLFMILYFTMAGSVWW VILTITWFLAAVFKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VREIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBTTWIQERC CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG TLANFVFCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CRUCKTAR STVTQCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG TLANFVFCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG TLANFVFCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG TLANFVFCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFFLLSN*CIWTOPQGYPQSSYG TLANFVFCSVRHGLALILQLCRFSIYTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFTLASITTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFFTLASITTOQMN LSIAIPAMVNNTAPPSQPNASTERPST CSRKGIVDEFTTLASITTOQMN CSRKGIVTOPGGYPQSSYG CSRKGIVDEFFTLASITTOQMN CSRKGIVTOPGGYPQSSYG CSRKGIVTOPGGYPQSSYG CSRKGIVTOPGGYPQSSYG CSRKGIVTOPGGYPQSSYG CSRKGI			}	1			
LLQKAIRNOK**CTVQQLSHCRLY\GEKTTAK] -	l	ł	j	ļ	ļ	
RSQREHYQQQSQEHGKWPDLKGPR			1		ĺ]	
1077		į .	1				
QYPALHRAGTEWQLSALHRAPRSTOPDKAC RIGYKAKQGYIIYRICVRRGGWKCPVPKAVT VYGKPVHHGVN*LKFAQSLQSVABEQ 1078 2428 A 8905 536 781 ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYPTLGLPTQSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MIFGVPWPEDMECSRFPDCDETYRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFIFVTFLIDVTRFYPERPIKCYAV WHMMVSLIFFUGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILITTWFLAAVFKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYTYEQAYRGIWBTTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFFLLSN*CIWIQPQGYPQSSYG TLANFVFVCSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1000	2427	1	8001	7.50		
RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT \YGKPVHHGVN*LKFAQSLQSVAEBQ 1078 2428 A 8905 536 781 ACPAENREVPEMAAGQAPHAGPGAGPGPA PALPFAATTGSRGQALCRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYPTLGLPTGSTFL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPLLNIA DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGYPWPEDMECSRFPDCDEFYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSSPCPNMYFRREELSFARYFIGLIS IICLSATLFTFVTFLIDVTRFYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNASIPAQYKA STVTQGSHNKACTMLFMILYFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFFLLSN*CIWTQPQGYPQSSYG TLANFVFCSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	10//	242/	A	9901	332	3	
1078 2428 A 8905 536 781 ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRQHI.HGPLH RP*QAAPALHAGCQLAPHPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIGFKPGQDKYFTLGLPTGSTTP+CYPKLI EYNKNGHI.SFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKLIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWEITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG CSRKGIVDEFFPLLSN*CLUVTQPQGYPQSSYG		ł	1	Į.	į	<u> </u>	
1078 2428 A 8905 536 781 ACPAENREVPEMAAGQAPHAGPGAGPGQPA PALPFAATPGSRGQALCRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHPPT NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTQSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLPSCEPTTLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEBTIEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMILFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VREIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWEITTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFFLLSN*CIWTQPQGYPQSSYG TLANFVP\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	l		1			}	
PALPFAATPGSRGQALČRGGRRRQHLHGPLH RP*QAAPALHAGCQLAPHIPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIGFKPGQDKYPTLGLFTQSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS HICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNAISIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSYG TLANPVFVCSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1078	2428	A	8905	536	781	
RP*QAAPALHAGCQLAPHPPT 1079 2429 A 8912 121 376 NLIWKLCVTERRLVILDNYDLASE/YEANKYI CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI EYNKNGHI SPKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNILNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFIFVTFLIDVTRFYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNAISIPAQYKA STVTQGSHNKACTMLFMILYFFIMAGSVWW VII.TITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VREIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWEITWIQERC CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVFVCSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST				1			
CNRIIQFKPGQDKYPTLGLPTGSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFL.CAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGF WCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFVGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBTTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST			ļ	1		•	
CNRIIQFKPGQDKYPTLGLPTGSTPL*CYPKLI EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFVGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1079	2429	A	8912	121	376 -	
EYNKNGHLSFKYVKTFSMDEY 1080 2430 A 8920 381 1788 SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLQQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFTFVFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	_]]				
GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GBPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	L	L					EYNKNGHLSFKYVKTFSMDEY
DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL YAPICMEYGRVTLPCRRLQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFITPVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFNGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVP\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1080	2430	A	8920	381	1788	
YAPICMEYGRVTLPCRRLCQRAYSECSKLME MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFNGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1				1	į	GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFF\GFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIGERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANPVP\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1		1		1		DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY SFLHVRDCSPPCPNMYFRREBLSFARYFIGLIS IICLSATLFITFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFF\GFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBTTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST			1			1	YAPICMEYGRVTLPCRRLCQRAYSECSKLME
SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS IICLSATLFIFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFFIGFILEDRVACNASIPAQYKA STVTQGSHNKACIMLFMILYFFIMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISGONLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1		1			1	
IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACIMLFMILYFFIMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST			l	1	1]	
WHMMVSLIFF\IGFLLEDRVACNA\SIPAQYKA STVTQGSHNKACTMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBTTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	ł		ł	}	ł)	
STVTQGSHNKACIMLFMILYFFTMAGSVWW VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWEITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1			}		1	
VILTITWFLAAVPKWGSEAIEKKALLFHASA WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST		1		1	ļ		
WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1	1	1				
VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWBITWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISGGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	l	1	1		İ	l	, ·
VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL VVIGCYFYEQAYRGIWETTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	į		1				§
VVIGCYFYEQAYRGIWETTWIQERC 1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1	ļ	1	j .	j	j	
1081 2431 A 8922 56 420 EERTKMSTGPDVKATVGDISSDGNLNVAQEE CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1					1	
CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1081	2431	A	8922	56	420	
TLANFVF\CSVRHGLALILQLCNFSIYTQQMN LSIAIPAMVNNTAPPSQPNASTERPST	1.701	~~~	**	1	~~	1 -20	CSRKGIVDEFFPLLSN*CIWTOPOGYPOSSYG
LSIAIPAMVNNTAPPSQPNASTERPST	1	1	Į.		l	1	
	Ì						
1082 2432 A 8923 355 1079 PFGTPSTMAVVKNKCLMKGGKKGVKKKVV	1082	2432	A	8923	355	1079	PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

to a second to

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Les pages 340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20